

A Simple DFT-aided Spatial Basis Expansion Model and Channel Estimation Strategy for TDD/FDD Massive MIMO Systems

Hongxiang Xie, Feifei Gao, Shun Zhang, and Shi Jin

Abstract

This paper proposes a new transmission strategy for the multiuser massive multiple-input multiple-output (MIMO) systems, including uplink/downlink channel estimation and user scheduling for data transmission. A discrete Fourier transform (DFT) aided spatial basis expansion model (SBEM) is first introduced to represent the uplink/downlink channels with much few parameter dimensions by exploiting angle reciprocity and the physical characteristics of the uniform linear array (ULA). With SBEM, both uplink and downlink channel estimation of multiusers can be carried out with very few amount of training resources, which significantly reduces the training overhead and feedback cost. Meanwhile, the pilot contamination problem in the uplink training is immediately relieved by exploiting the spatial information of users. To enhance the spectral efficiency and to fully utilize the spatial resources, we also design a greedy user scheduling scheme during the data transmission period. Compared to existing low-rank models, the newly proposed SBEM offers an alternative for channel acquisition without need of channel statistics for both TDD and FDD systems based on the angle reciprocity. Moreover, the proposed method can be efficiently deployed by the fast Fourier transform (FFT). Various numerical results are provided to corroborate the proposed studies.

Index Terms

Massive MIMO, spatial basis expansion model (SBEM), discrete Fourier transform (DFT), direction of arrival (DOA), angle reciprocity.

H. Xie and F. Gao are with Tsinghua National Laboratory for Information Science and Technology (TNList) Beijing 100084, P. R. China (e-mail: xiehx14@mails.tsinghua.edu.cn, feifeigao@ieee.org).

S. Zhang is with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, P. R. China (email: zhangshunsdu@gmail.com).

S. Jin is with the National Communications Research Laboratory, Southeast University, Nanjing 210096, P. R. China (email: jinshi@seu.edu.cn).

I. INTRODUCTION

Large-scale multiple-input multiple-output (MIMO) or “massive MIMO” systems [1] have drawn considerable interest from both academia and industry. Theoretically, massive MIMO systems can almost perfectly relieve the inter-user interference in multiuser MIMO (MU-MIMO) systems with simple linear transceivers [2]. It was shown in [3] that each antenna element of a very large MIMO system consumes exceedingly low power, and the total power can be made inversely proportional to the number of antennas. Other advantages, such as high spectral efficiency, security, robustness or reliable linkage, also play key roles in promoting massive MIMO systems more appealing for the next generation of wireless systems [3], [4].

However, all these potential gains of massive MIMO systems rely heavily on the perfect channel state information (CSI) at the base station (BS). From the conventional orthogonal training strategy [5], the optimal number of training streams should be the same as the number of the transmit antennas and the length of the training should be no less than the number of transmit antennas. Hence, downlink training in massive MIMO system requires huge number of orthogonal training sequences. This severe overhead as well as the accompanied high calculation complexity will overwhelm the system performance and mitigate any possible improvement. For uplink training, the conventional channel estimation methods are generally feasible. However, as the number of users or the number of user’s antennas grows, the increased pilot overhead will deteriorate the system efficiency and becomes the system bottleneck. If the non-orthogonal sequences are used for uplink training, then the so caused pilot contamination will also deteriorate the system performance.

For time division duplexing (TDD) massive MIMO systems, downlink CSI can be obtained by leveraging the channel reciprocity [6], which has promoted quite many research works [7]–[12]. However, in practice the reciprocity between uplink and downlink may not exactly hold even for TDD systems due to the calibration error between the downlink/uplink RF chains [13]. In addition, the property of channel reciprocity has been proven to be robust only for the single-cell scenario [14], and this will undoubtedly increase the pressure of multi-cell coordination. On the other side, the downlink channel estimation for frequency division duplexing (FDD)

massive MIMO system is always deemed as a difficult problem since the channel reciprocity does not hold and cannot be used to simplify the estimation. The authors of [2], [15], [16] applied the closed-loop training schemes to sequentially design the optimal pilot beam patterns. The compressive sensing (CS) based feedback-reduction in [17] and the distributed compressive channel estimation in [18] exploited the channel statistics to reduce the heavy burden of feeding back large amount of measurements.

Except for the above regular attempts, a new way to design the transmission strategy for massive MIMO system is to exploit the low-rank approximation of channel covariance matrix [19]–[21]. Based on the assumption that the angular spread (AS) of the incident signals at BS from each user is narrow and the antenna elements only have half-wave length spacing, the authors in [19]–[21] proposed to reduce the dimensionality of the effective channels through eigen-decomposition of channel covariance matrices. All these covariance-aware methods could be categorized into *spatial division multiplexing* that utilizes non-overlapping spatial information of different users to realize the orthogonal transmission. Meanwhile, compressed channel sensing has been widely adopted for channel sparsity models, such as the Karhunen-Loeve transform basis representation in [18] and the virtual channel representation in [22]. As the channel sparsity patterns in these existing methods are often assumed unknown, high-complexity nonlinear CS reconstruction procedures are thus inevitable.

In this paper, we propose an alternative low-rank model based on the discrete Fourier transform (DFT) of the steering vectors for the uniform linear array (ULA). The new model could exploit the spatial information of the users and is then named as spatial basis expansion model (SBEM). It is shown that the uplink/downlink channel estimation of multiusers can be carried out with very few training resources, and thus the overhead of training and feedback can also be reduced significantly. Meanwhile, the pilot contamination in uplink training can be immediately relieved. To enhance the spectral efficiency during the data transmission, we propose a greedy user scheduling algorithm where users with orthogonal spatial information are allowed to transmit simultaneously. Compared to the existing low-rank models [18]–[20], [22], the proposed one offers following several benefits: (1) SBEM can exploit the angle reciprocity and simplify the DL

training from the UL channel estimation for both TDD/FDD massive MIMO systems; (2) SBEM does not need the knowledge of channel covariance; (3) SBEM could be simply implemented by the fast Fourier transform (FFT) and the partial FFT [23]. By contrast, [19] and [20] need EVD for high-dimensional covariance matrices, while [18] and [22] require nonlinear optimization. Various numerical results are provided to evaluate performances of the proposed method.

The rest of the paper is organized as follows. In section II, the system model as well as the channel characteristics with narrow incident signals are described. The SBEM aided channel estimations for uplink/downlink transmission are presented in section III. Section IV designs the user scheduling for data transmission, followed by simulations in section V. Finally, conclusions are drawn in section VI.

Notations: Vectors and matrices are boldface small and capital letters; the transpose, complex conjugate, Hermitian, inverse, and pseudo-inverse of the matrix \mathbf{A} are denoted by \mathbf{A}^T , \mathbf{A}^* , \mathbf{A}^H , \mathbf{A}^{-1} and \mathbf{A}^\dagger , respectively; $\text{tr}(\mathbf{A})$ is the trace of \mathbf{A} ; $[\mathbf{A}]_{ij}$ is the (i, j) th entry of \mathbf{A} ; the entry index of vector and matrix starts from 0; $\text{diag}\{\mathbf{a}\}$ denotes a diagonal matrix with the diagonal elements constructed from \mathbf{a} , while $\text{diag}\{\mathbf{A}\}$ denotes a vector whose elements are extracted from the diagonal components of \mathbf{A} ; \mathbf{I} is the identity matrix with appropriate size, and $\mathbb{E}\{\cdot\}$ is the statistical expectation; $\lceil x \rceil$ denotes the smallest integer no less than x , $\lfloor x \rfloor$ denotes the largest integer no more than x and $\lceil x \rceil$ denotes the integer closest to x ; $[\mathbf{H}]_{:, \mathcal{D}}$ denotes the sub-matrix of \mathbf{H} by collecting the columns indexed by \mathcal{D} , and $[\mathbf{H}]_{\mathcal{D}, :}$ denotes the sub-matrix of \mathbf{H} by collecting the rows indexed by \mathcal{D} ; $[\mathbf{h}]_{\mathcal{D}, :}$ indicates the sub-vector of \mathbf{h} by keeping the elements indexed by \mathcal{D} ; $|\mathcal{D}|$ denotes the cardinality of the set \mathcal{D} ; and $\|\mathbf{h}\|$ denotes the Euclidean norm of \mathbf{h} ; “ \setminus ” defines the set subtraction operation.

II. SYSTEM MODEL AND CHANNEL CHARACTERISTICS

A. System Model

Array structure based physical channel models have been widely adopted for MIMO systems, such as the spatial channel model (SCM) [26], which exploits the array manifold and information of the direction of arrivals (DOA) as well as the direction of departures (DOD) of propagation signals. For massive MIMO systems, the significantly improved spatial resolution of large-scale

antenna arrays as well as plenty of emerging low-complexity DOA acquisition techniques [27]–[30] have further promoted these physical channel models [19]–[21]. Many works [13], [21], [31] of massive MIMO systems require BS to be elevated at a very high altitude, say on the top of a high building or a dedicated tower such that there are few surrounding scatterers at the end of BS. In this case, the incident angular spread seen by the BS array is usually limited in a narrow region [19]–[21].

Let us consider a multiuser massive MIMO system, where BS is equipped with $M \gg 1$ antennas in the form of ULA, and K single-antenna users are randomly distributed in the coverage area. For better illustration of our key idea, we assume the channel is flat fading for the time being as did in [19], [20].

Consider the classical “one-ring model” adopted in [20], where user- k located at D_k meters away from BS is surrounded by a ring of $P \gg 1$ local scatterers (see Fig. 1) with the radius R_k . Then the propagation from user- k to BS is composed of P rays and the corresponding $M \times 1$ uplink channel can be expressed as [19]:

$$\mathbf{h}_k = \frac{1}{\sqrt{P}} \sum_{p=1}^P \alpha_{kp} \mathbf{a}(\theta_{kp}), \quad (1)$$

where $\alpha_{kp} \sim \mathcal{CN}(0, \xi_{kp})$ represents the complex gain of the p -th ray and is independent and identically distributed (i.i.d.) from each other. Moreover, $\mathbf{a}(\theta_{kp}) \in \mathbb{C}^{M \times 1}$ is the array manifold vector and has the form

$$\mathbf{a}(\theta_{kp}) = \left[1, e^{j \frac{2\pi d}{\lambda} \sin \theta_{kp}}, \dots, e^{j \frac{2\pi d}{\lambda} (M-1) \sin \theta_{kp}} \right]^T, \quad (2)$$

where d is the antenna spacing, λ denotes the signal carrier wavelength, and θ_{kp} represents the DOA of the p -th ray.

Since the distance D_k from BS to user- k is always much larger than the radius R_k , the incident rays will be constrained within a narrow angular spread (AS) $\Delta\theta_k \approx \arctan(R_k/D_k)$ [20]. In other words, the incident angles of user- k with mean DOA θ_k is limited in the narrow angular range $[\theta_k - \Delta\theta_k, \theta_k + \Delta\theta_k]$. Within this narrow angular range, it is obvious that $\mathbf{a}(\theta_{kp})$'s are highly correlated with each other such that the channel covariance matrix of user- k , defined as $\mathbf{R}_{h_k} = \mathbb{E}\{\mathbf{h}_k \mathbf{h}_k^H\}$ approximately possesses the low-rank property. Based on this low-rank property, [18],

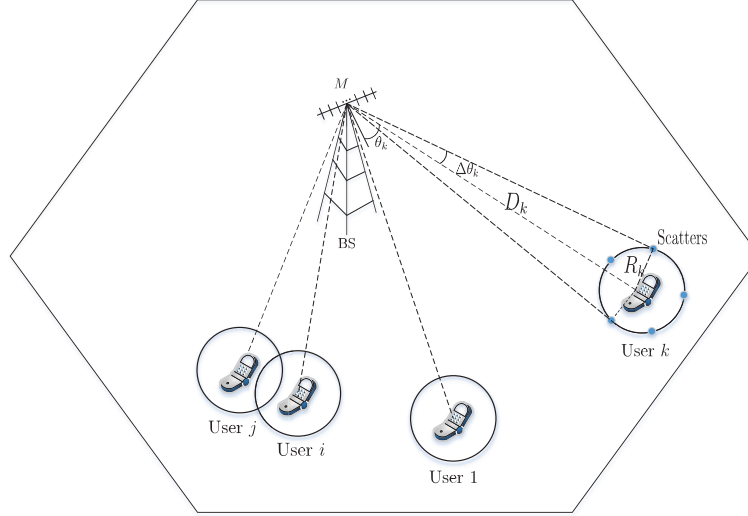


Fig. 1. System of one-ring model. Users are randomly distributed and surrounded by P local scatterers. The mean DOA and AS of user- k are θ_k and $\Delta\theta_k$, respectively.

[19], [20] and [22] presented low-rank channel models and designed the corresponding channel estimation to reduce the channel overhead as well as computational complexity. However, the methods in [19] and [20] need eigen-decomposition for an $M \times M$ channel covariance matrix, while [18] and [22] demand for nonlinear optimization, making their methods still complex for practical implementations.

In this paper, we present an alternative way to exploit the such a low-rank property, targeting at further reducing the channel estimation complexity. Due to the high correlations among $\mathbf{a}(\theta_{kp})$, $p = 1, \dots, P$, \mathbf{h}_k can be expanded from some orthogonal basis as

$$\mathbf{h}_k = \sum_{q=1}^{\nu} \kappa_q \mathbf{b}_q, \text{ for } k = 1, \dots, K, \quad (3)$$

where \mathbf{b}_q 's are basis vectors to be determined and κ_q 's are the corresponding coefficients. Equation (3) is also known as basis expansion model (BEM) as long as we could find a set of uniform basis vectors \mathbf{b}_q 's for any possible \mathbf{h}_k [32]. Then the task of channel estimation will be greatly simplified and will be converted to estimating ν constants κ_q 's only.

B. Characteristics of ULA and Channel Vectors

Define the normalized DFT of the channel vector \mathbf{h}_k as $\tilde{\mathbf{h}}_k = \mathbf{F}\mathbf{h}_k$, where \mathbf{F} is the $M \times M$ DFT matrix whose (p, q) th element is $[\mathbf{F}]_{pq} = e^{-j\frac{2\pi}{M}pq}/\sqrt{M}$.

Property 1: For the simplest 1-ray case, i.e., $\mathbf{h}_k = \alpha_k \mathbf{a}(\theta_k)$, $\tilde{\mathbf{h}}_k$ is approximately a sparse vector that contains the spatial information of the impinging signal.

Proof: The q -th component of $\tilde{\mathbf{h}}_k$ is computed as

$$\begin{aligned} [\tilde{\mathbf{h}}_k]_q &= \frac{\alpha_k}{\sqrt{M}} \sum_{m=0}^{M-1} e^{-j(\frac{2\pi}{M}mq - \frac{2\pi}{\lambda}md \sin \theta_k)} \\ &= \frac{\alpha_k}{\sqrt{M}} e^{-j\frac{M-1}{2}[\frac{2\pi}{M}q - \frac{2\pi}{\lambda}d \sin \theta_k]} \cdot \frac{\sin[(\frac{2\pi}{M}q - \frac{2\pi}{\lambda}d \sin \theta_k) \frac{M}{2}]}{\sin[(\frac{2\pi}{M}q - \frac{2\pi}{\lambda}d \sin \theta_k) \frac{1}{2}]}. \end{aligned} \quad (4)$$

When $M\frac{d}{\lambda}\sin\theta_k$ equals to some integer q_0 , then $\tilde{\mathbf{h}}$ only has one non-zero element $[\tilde{\mathbf{h}}_k]_{q_0} = \alpha_k\sqrt{M}$. In this specific case, $\tilde{\mathbf{h}}_k$ is highly sparse and all powers are concentrated on the q_0 -th DFT point, as illustrated in Fig. 2(a). However, for most other cases, $M\frac{d}{\lambda}\sin\theta_k$ is not an integer, and the channel power will leak from the $(\lfloor M\frac{d}{\lambda}\sin\theta_k \rfloor)$ -th DFT point to other DFT points, as shown in Fig. 2(b). In fact, DFT outputs are discrete samples of discrete-time Fourier transform (DTFT) of $\mathbf{a}(\theta_k)$, i.e., a Sinc function, at the points of $\frac{2\pi q}{M}$, $q = 0, \dots, M-1$. It is then easily known that the degree of leakage is inversely proportional to M but is proportional to the deviation $(M\frac{d}{\lambda}\sin\theta_k - \lfloor M\frac{d}{\lambda}\sin\theta_k \rfloor)$. Hence when M is sufficiently large, $\tilde{\mathbf{h}}_k$ can still be approximated by a sparse vector with most of power being concentrated around $M\frac{d}{\lambda}\sin\theta_k$. Specifically, for the ideal case $M \rightarrow \infty$, there always exists a q_0 that satisfies $q_0 = M\frac{d}{\lambda}\sin\theta_k$ for any possible θ_k and then the power leakage is eliminated. ■

From *Property 1*, it is clear that most channel power is concentrated on DFT points in the vicinity of $\lfloor M\frac{d}{\lambda}\sin\theta_k \rfloor$ when M is large. It is then of interest to find the minimal number of DFT points around $\lfloor M\frac{d}{\lambda}\sin\theta_k \rfloor$ that contain at least η percentage of the total power. Define the corresponding set of DFT points as \mathcal{C}_k , which can be found from the following optimization:

$$\begin{aligned} \min_{\mathcal{C}_k} \quad & C_k \triangleq |\mathcal{C}_k| \\ \text{s.t.} \quad & \sum_{q \in \mathcal{C}_k} \frac{1}{M} \left\{ \frac{\sin[(\frac{2\pi}{M}q - \frac{2\pi}{\lambda}d \sin \theta_k) \frac{M}{2}]}{\sin[(\frac{2\pi}{M}q - \frac{2\pi}{\lambda}d \sin \theta_k) \frac{1}{2}]} \right\}^2 \geq \eta. \end{aligned} \quad (5)$$

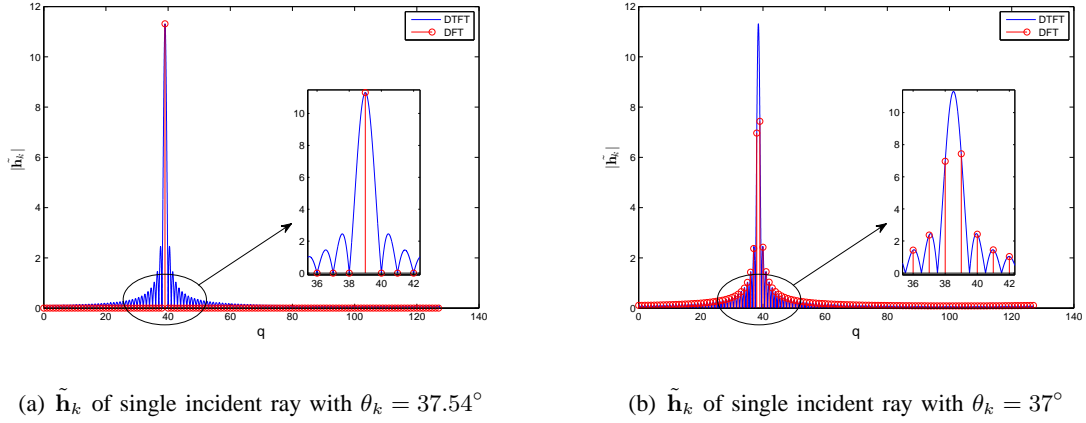


Fig. 2. Examples of $\tilde{\mathbf{h}}_k$ for 1-ray case, where $d = \frac{\lambda}{2}$ and $M = 128$.

TABLE I

AN EXAMPLE OF OFF-LINE TABLE OF C_k , WHERE $M = 128$, $\eta = 0.95$ AND $\theta_k \in [1^\circ, 90^\circ]$

θ_k	1°	3°	5°	7°	9°	11°	13°	15°	17°	19°	21°	23°	25°	27°	29°
C_k	1	8	9	4	1	5	9	9	7	4	1	1	1	1	1
θ_k	31°	33°	35°	37°	39°	41°	43°	45°	47°	49°	51°	53°	55°	57°	59°
C_k	1	3	7	10	6	1	8	6	4	7	6	1	9	7	3
θ_k	61°	63°	65°	67°	69°	71°	73°	75°	77°	79°	81°	83°	85°	87°	89°
C_k	1	1	1	1	6	10	4	4	8	4	5	9	5	1	1

The optimization in (5) is difficult to solve in closed-form. Nevertheless, since (5) is merely dependent on the array configuration, we can establish the off-line tables for $\theta_k \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and some pre-defined η 's. An example of the off-line table of C_k for a single incident ray is given in Tab. I with $M = 128$ and $\eta = 95\%$, where C_k is a function θ_k and η . It is seen from Tab. I that the maximal number of leakage points that contain at least 95% of the channel power is 10 for any incident angle, which is very small compared to the total antenna number $M = 128$. Hence, we could safely treat $\tilde{\mathbf{h}}_k$ as a sparse vector for massive antenna system.

Property 2: For the multi-ray case (1), let us define \mathcal{D}_k as the index set of the continuous DFT points that contain at least η percentage of the total channel power. The upper bound of

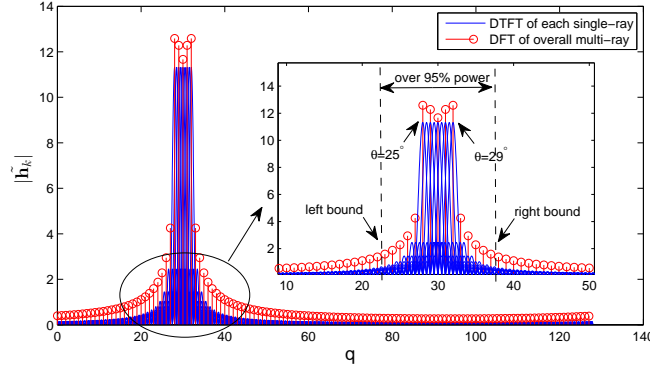


Fig. 3. Example of multi-ray channel with incident angles in $[25^\circ, 29^\circ]$. DTFT of each single-ray channel (blue lines) as well as DFT of the overall multi-ray channel (red points) are depicted, respectively.

the cardinality of \mathcal{D}_k can be expressed as

$$|\mathcal{D}_k| \leq \lceil 2M \frac{d}{\lambda} \cdot |\cos \theta_k| \cdot \Delta \theta_k + 1 \rceil + C_{\max}, \quad (6)$$

where $C_{\max} \triangleq \max C_k$ for given η and all $\theta \in [\theta_k - \Delta \theta_k, \theta_k + \Delta \theta_k]$.

Proof: The left bound of \mathcal{D}_k is determined by the DFT of the leftmost ray with $\theta_{kp} = \theta_k - \Delta \theta_k$ and can be expressed as $\lfloor M \frac{d}{\lambda} \sin(\theta_k - \Delta \theta_k) \rfloor - \lceil C_{\max}/2 \rceil$ according to *Property 1*. Similarly, the right bound of \mathcal{D}_k depends on the DFT of the rightmost ray with $\theta_{kp} = \theta_k + \Delta \theta_k$ and can be expressed as $\lceil M \frac{d}{\lambda} \sin(\theta_k + \Delta \theta_k) \rceil + \lceil C_{\max}/2 \rceil$. Then for any single ray with incident DOA inside $[\theta_k - \Delta \theta_k, \theta_k + \Delta \theta_k]$, the corresponding main leakage points that count for η percentage of its own power will be included in \mathcal{D}_k . From the above discussion, the upper bound of the cardinality of \mathcal{D}_k can be directly obtained as

$$\begin{aligned} |\mathcal{D}_k| &\leq \lceil M \frac{d}{\lambda} \sin(\theta_k + \Delta \theta_k) \rceil - \lfloor M \frac{d}{\lambda} \sin(\theta_k - \Delta \theta_k) \rfloor + 1 + C_{\max} \\ &\leq \lceil 2M \frac{d}{\lambda} \cdot |\cos \theta_k| \cdot \Delta \theta_k + 1 \rceil + C_{\max}. \end{aligned} \quad (7)$$

■

An example of a 9-ray channel with incident angles in $[25^\circ, 29^\circ]$ is given in Fig. 3, where the DTFT of each single ray as well as the DFT of the overall multi-rays are depicted respectively. The numerical results in this example show that the actual cardinality of \mathcal{D}_k containing 95% power is 15, and the upper bound calculated from (7) is also 15.

Following *Property 2* and bearing in mind that $\Delta\theta_k$ is small, it is readily known that $|\mathcal{D}_k| \leq \lceil \frac{2Md}{\lambda} \cdot |\cos \theta_k| \cdot \Delta\theta_k + 1 \rceil + C_{\max}$ is still small compared to M . Hence, $\tilde{\mathbf{h}}_k$ corresponding to the multi-ray case is also approximately sparse with most power being contained in limited number of entries.

Remark 1: The sparsity of mutli-ray case is very obviously if $M \rightarrow \infty$ is assumed, as did in most massive MIMO works [19], [21]. Nevertheless, we do not make such an ideal assumption throughout this paper in order to present a practical solution for large but limited number of array antennas.

Therefore, the key idea of this paper that is to approximate the channel vector with fewer parameters as

$$\mathbf{h}_k = \mathbf{F}^H \tilde{\mathbf{h}}_k \approx [\mathbf{F}^H]_{:, \mathcal{D}_k} \left[\tilde{\mathbf{h}}_k \right]_{\mathcal{D}_k, :} = \sum_{q \in \mathcal{D}_k} \tilde{h}_{k,q} \mathbf{f}_q, \quad (8)$$

where $\tilde{h}_{k,q} \triangleq [\tilde{\mathbf{h}}_k]_q$ denotes the q -th element of $\tilde{\mathbf{h}}_k$ while \mathbf{f}_q is the q -th column of \mathbf{F}^H . Comparing with (3), the expansion in (8) is also in the form of BEM where the basis vectors $\mathbf{b}_q \triangleq \mathbf{f}_q$ are orthogonal to each other. Hence, we only need to estimate the limited BEM parameters $\tilde{h}_{k,q}$.

Interestingly, the DFT vector \mathbf{f}_q coincides with the steering vector as $\mathbf{f}_q = \mathbf{a}(\theta_q)$ where $\theta_q = \arcsin \frac{q\lambda}{Md}$, which means that \mathbf{f}_q formulates an array beam towards the physical direction $\theta_q = \arcsin \frac{q\lambda}{Md}$. Hence, all beams \mathbf{f}_q , $q \in \mathcal{D}_k$ will point towards the AS of user- k and are orthogonal to each other. Consequently, the beam indices \mathcal{D}_k can be viewed as the *spatial signature* of user- k [33], and (8) can be deemed as the *spatial BEM* (SBEM), as named after the popular temporal BEM [32].

Property 3: Define

$$\Phi(\phi) = \text{diag} \left\{ [1, e^{j\phi}, \dots, e^{j(M-1)\phi}] \right\}, \quad (9)$$

where $\phi \in [-\frac{\pi}{M}, \frac{\pi}{M}]$ is a *shift parameter*. The operation $\tilde{\mathbf{h}}_k^{\text{ro}} = \mathbf{F}\Phi(\phi)\mathbf{h}_k$ can further concentrate channel power within fewer entries of $\tilde{\mathbf{h}}_k^{\text{ro}}$ for certain value of ϕ , and this operation is named as *spatial rotation*.

Proof: Consider the 1-ray case for the ease of illustration. In Fig. 2(b), when the DOA of the incident signal is not $\arcsin \frac{q_0\lambda}{Md}$ for some integer q_0 , i.e., mismatched with DFT points, then power leakage will happen. Formulate a new channel vector as $\tilde{\mathbf{h}}_k^{\text{ro}} = \mathbf{F}\Phi(\phi)\mathbf{h}_k$. According to

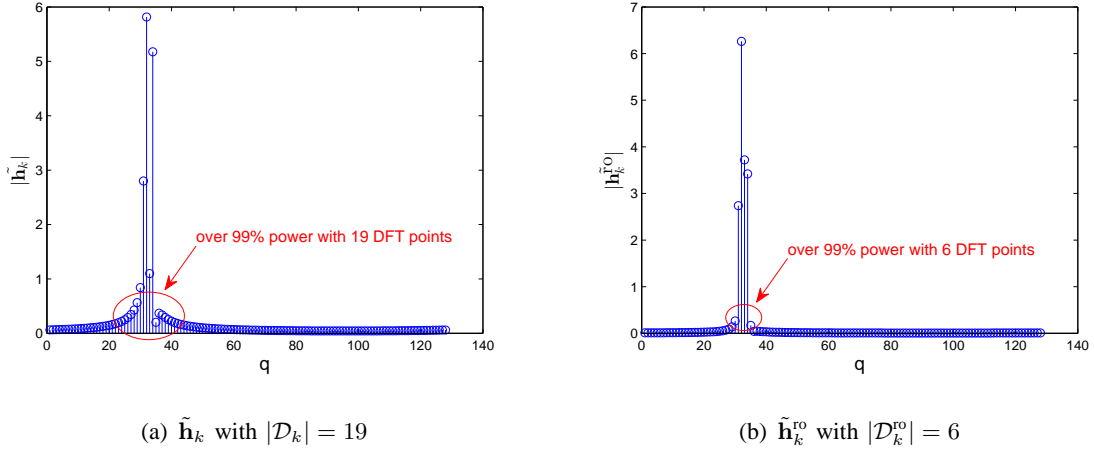


Fig. 4. Comparison of multi-ray channels with/without spatial rotation, where $\theta_k = 30^\circ$, $\Delta\theta_k = 2^\circ$, $M = 128$, $d = \frac{\lambda}{2}$, $\eta = 0.99$ and the optimal $\phi_k = -0.0142$ (in radian), within $[-\frac{\pi}{128}, \frac{\pi}{128}]$.

the analysis in *Property 1*, let us gradually change ϕ from $-\frac{\pi}{M}$ to $\frac{\pi}{M}$.¹ Then, there will exist a $\phi_k = (\frac{2\pi q_0}{M} - \frac{2\pi d}{\lambda} \sin \theta_k)$ that makes $\tilde{\mathbf{h}}_k^{\text{ro}}$ possess only one non-zero element $\tilde{h}_{k,q_0}^{\text{ro}}$ at q_0 . For example in Fig. 2, spatial rotation with $\phi_k = 0.024$ radian can help to strengthen the sparsity of $\tilde{\mathbf{h}}_k$ in Fig. 2(b) to the form in Fig. 2(a).

For the multi-ray cases, we can also formulate $\tilde{\mathbf{h}}_k^{\text{ro}} = \mathbf{F}\Phi(\phi)\mathbf{h}_k$, and define $\mathcal{D}_k^{\text{ro}}$ as the continuous index set such that $[\tilde{\mathbf{h}}_k^{\text{ro}}]_{\mathcal{D}_k^{\text{ro}},:}$ contains at least η percentage of the channel power. Next, we can search ϕ from $-\frac{\pi}{M}$ to $\frac{\pi}{M}$ and select the optimal ϕ_k that minimizes $|\mathcal{D}_k^{\text{ro}}|$. An example of multi-ray channel with $\theta_k = 30^\circ$ and $\Delta\theta_k = 2^\circ$ is given in Fig. 4, where $M = 128$, $d = \frac{\lambda}{2}$ and $\eta = 0.99$. It can be seen that the cardinality of $\mathcal{D}_k^{\text{ro}}$ is only $|\mathcal{D}_k^{\text{ro}}| = 6$ after spatial rotation, while the cardinality before the rotation is $|\mathcal{D}_k| = 19$. ■

Based on the *Property 3* and similar to (8), there is

$$\mathbf{h}_k = \Phi(\phi_k)^H \mathbf{F}^H \tilde{\mathbf{h}}_k^{\text{ro}} \approx \Phi(\phi_k)^H [\mathbf{F}^H]_{:, \mathcal{D}_k^{\text{ro}}} [\tilde{\mathbf{h}}_k^{\text{ro}}]_{\mathcal{D}_k^{\text{ro}},:} = \sum_{q \in \mathcal{D}_k^{\text{ro}}} \tilde{h}_{k,q}^{\text{ro}} \Phi(\phi_k)^H \mathbf{f}_q, \quad (10)$$

where $\Phi(\phi_k)^H \mathbf{f}_q$, $q \in \mathcal{D}_k^{\text{ro}}$ are the new equivalent basis vectors. Interestingly, these new basis vectors are still mutually orthogonal, and thus (10) is also a type of SBEM. In many cases,

¹The reason for $\phi_k \in [-\frac{\pi}{M}, \frac{\pi}{M}]$ lies in that the resolution of DFT is $\frac{2\pi}{M}$.

$|\mathcal{D}_k^{\text{ro}}|$ is much less than $|\mathcal{D}_k|$ after the optimal rotation, and hence the number of the channel parameters to be estimated is further reduced.

Remark 2: In fact, the operation $\Phi(\phi_k)\mathbf{h}_k$ can be viewed as rotating the orthogonal beams \mathbf{f}_q 's by the same angle such that the new beams $\Phi(\phi_k)^H\mathbf{f}_q$'s point towards user- k more accurately while still keeping their orthogonality.

Property 4: Define a set $\mathcal{E}_k^{\text{ro}}$ that is sufficiently far from $\mathcal{D}_k^{\text{ro}}$. Then, $[\mathbf{F}\Phi(\phi)\mathbf{h}_k]_{\mathcal{E}_k^{\text{ro}},:}$ always has ignorable value for any $\phi \in [-\pi/M, \pi/M]$ as compared with $[\mathbf{F}\Phi(\phi)\mathbf{h}_k]_{\mathcal{D}_k^{\text{ro}},:}$. The reason is that $\mathbf{F}\Phi(\phi)\mathbf{h}_k$ is the sampling of the DTFT of \mathbf{h}_k after being shifted by a small value ϕ . Hence, as shown in Fig. 4, $[\mathbf{F}\Phi(\phi)\mathbf{h}_k]_{\mathcal{E}_k^{\text{ro}},:}$ still possess very small value compared to the sampling points in $\mathcal{D}_k^{\text{ro}}$. This property will be used later during the channel estimation.

Nevertheless, the practical standards normally regulate a fixed number of channel parameters to be estimated instead of considering a dynamic number $|\mathcal{D}_k^{\text{ro}}|$. Denote τ as the number of the channel parameters that the system could handle, and define the set containing continuous τ integers as $\mathcal{B}_k^{\text{ro}}$. Then we should select $\mathcal{B}_k^{\text{ro}}$ as well as the shift parameter ϕ_k such that $[\tilde{\mathbf{h}}_k^{\text{ro}}]_{\mathcal{B}_k^{\text{ro}},:}$ possesses the maximum channel power, i.e.,

$$\max_{\phi_k, \mathcal{B}_k^{\text{ro}}} \left\| [\tilde{\mathbf{h}}_k^{\text{ro}}]_{\mathcal{B}_k^{\text{ro}},:} \right\|^2 \quad \text{subject to } |\mathcal{B}_k^{\text{ro}}| = \tau. \quad (11)$$

The above optimization can be achieved simply by sliding window of size τ over the elements in $\tilde{\mathbf{h}}_k^{\text{ro}}$ together by a one dimensional search over $\phi \in [-\pi/M, \pi/M]$.

III. CHANNEL ESTIMATION WITH SBEM

Assume the current cell is allocated $\tau < K$ orthogonal training sequences of length $L < T$, where T is the channel coherence interval, for both uplink and downlink training. Denote the corresponding orthogonal training set in the considered cell as $\mathcal{S}_{\text{cell}} = \{\mathbf{s}_1, \dots, \mathbf{s}_\tau\}$ with $\mathbf{s}_i^H \mathbf{s}_j = L\sigma_p^2 \cdot \delta(i - j)$, where σ_p^2 is the pilot signal training power.

We propose a new uplink/downlink transmission framework that utilizes the spatial signatures to realize the orthogonal training and data transmission among different users. As shown in Fig. 5, the transmissions between BS and users always start from an uplink preamble to obtain the spatial signature of each user. Then users are grouped and scheduled for uplink/downlink training

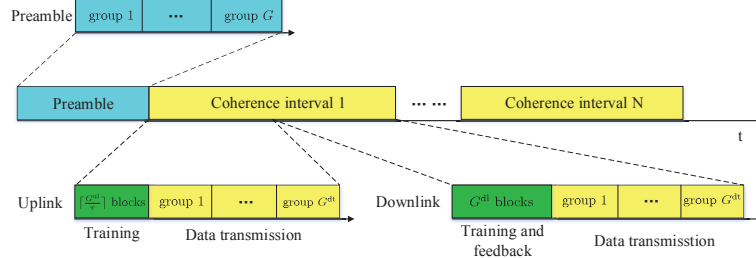


Fig. 5. The communication process of the new framework. *preamble* is used to collect the information about DOAs, then followed by uplink/downlink training and data transmission.

and data transmission based on their spatial signatures such that the orthogonal transmission is achieved.

It is worth mentioning that based on the proposed SBEM, once the spatial signatures of users are obtained in the preamble, the reduced-dimensional channels can be estimated through traditional linear LS method. However, to make the proposed strategy complete and address some ideas in detail, we will generally show the whole procedures of proposed channel estimation.

A. Obtain Spatial Information through Uplink Preamble

For the ease of illustration, assume $K = G\tau$ where $G \geq 1$ is an integer. Following most standards, there exists a relative long uplink training period called *preamble* at the very beginning of any transmission, while the functionality of the preamble in the proposed framework is mainly to obtain the spatial signatures of different users. Since we do not assume any prior spatial information, we will have to divide K users into G groups, each containing τ users such that τ orthogonal training sequences is sufficient for each group. Then, the conventional uplink channel estimation will be applied for each group during the preamble period, and hence the length of the preamble is GL .

Taking the first group as an example, the received training signals at BS is given by

$$\mathbf{Y} = \mathbf{H}\mathbf{D}^{1/2}\mathbf{S}^H + \mathbf{N} = \sum_{i=1}^{\tau} \sqrt{d_i} \mathbf{h}_i \mathbf{s}_i^H + \mathbf{N}, \quad (12)$$

where $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_\tau] \in \mathbb{C}^{M \times \tau}$ is the uplink channel matrix from the first τ users; $\mathbf{S} \triangleq [\mathbf{s}_1, \dots, \mathbf{s}_\tau] \in \mathbb{C}^{L \times \tau}$ contains τ training sequences assigned to this cell; $\mathbf{D} \triangleq \text{diag}\{d_1, \dots, d_\tau\}$

and $d_k \triangleq \frac{P_k^{\text{ut}}}{L\sigma_p^2}$ is used to satisfy the uplink training energy constraint P_k^{ut} of user- k ; \mathbf{N} is the independent additive white Gaussian noise matrix with elements distributed as i.i.d $\mathcal{CN}(0, \sigma_n^2)$.

Then \mathbf{h}_k can be estimated through LS as

$$\hat{\mathbf{h}}_k = \frac{1}{\sqrt{d_k}L\sigma_p^2} \mathbf{Y} \mathbf{s}_k = \mathbf{h}_k + \frac{1}{\sqrt{P_k^{\text{ut}}/\sigma_n^2}} \mathbf{n}_k, \quad (13)$$

where $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ denotes the normalized Gaussian white noise vector.

Repeating the similar operations in (13) for all G groups yields the channel estimates for all K users. The next step is to obtain the optimal rotation ϕ_k and extract the spatial signature $\mathcal{B}_k^{\text{ro}}$ of size τ that contains the maximum power of $\mathbf{F}\Phi(\phi_k)\hat{\mathbf{h}}_k$, as described in (11). Then, the channel vector can be approximated by

$$\mathbf{h}_k = \Phi(\phi_k)^H \mathbf{F}^H \tilde{\mathbf{h}}_k^{\text{ro}} \approx \Phi(\phi_k)^H [\mathbf{F}^H]_{:, \mathcal{B}_k^{\text{ro}}} \left[\tilde{\mathbf{h}}_k^{\text{ro}} \right]_{\mathcal{B}_k^{\text{ro}}, :}. \quad (14)$$

Remark 3: Contrast to existing works [19], [20] where channel covariances are directly assumed to be known, we here design a concrete method to acquire the spatial information for all users.

B. Uplink Training with User Grouping

Channel information obtained from preamble may only last for a short period, say one coherent time, while it should be tracked or estimated again in the later transmission. Since the channel coherence time is in the level of millisecond while a user and its surrounding obstacles may not physically change its position in the comparable time, we may treat DOA and AS of a user as unchanged within several or even tens of the channel coherence times. Hence, the spatial signatures $\mathcal{B}_k^{\text{ro}}$ of each user could be deemed as unchanged, and only the accompanied $\left[\tilde{\mathbf{h}}_k^{\text{ro}} \right]_{\mathcal{B}_k^{\text{ro}}, :}$ should be re-estimated in the later transmission. Moreover, the non-overlapping properties of different spatial signatures could also be utilized to release the pressure of insufficient orthogonal training sequences.

Let us then divide users into different groups according to their spatial signatures. Specifically, users are allocated to the same group if their spatial signatures do not overlap and are separated by a certain guard interval Ω , i.e.,

$$\mathcal{B}_k^{\text{ro}} \cap \mathcal{B}_l^{\text{ro}} = \emptyset, \quad \text{and} \quad \text{dist}(\mathcal{B}_k^{\text{ro}}, \mathcal{B}_l^{\text{ro}}) \geq \Omega, \quad (15)$$

where $\text{dist}(\mathcal{B}_1, \mathcal{B}_2) \triangleq \min |b_1 - b_2|$, $\forall b_1 \in \mathcal{B}_1, \forall b_2 \in \mathcal{B}_2$, and the value of Ω depends on the tolerance of users for the interference due to pilot reusing.

Assume that all users are divided into G^{ut} groups based on (15) and denote the user index set of the g -th group as $\mathcal{U}_g^{\text{ut}}$. We expect that $G^{\text{ut}} \ll K$ since $\tau = |\mathcal{B}_k^{\text{ro}}|$ is much smaller than M and users are randomly distributed in the service region.

As channels of different users in the same group could be discriminated by their spatial signatures, we could assign the same training sequence to all users in one group. Let us first consider the case $G^{\text{ut}} \leq \tau$ and allocate \mathbf{s}_i to the i -th group. Then, all K users in G^{ut} groups send their training sequences simultaneously, and the received signals at BS can be expressed as

$$\mathbf{Y} = \sum_{i=1}^{G^{\text{ut}}} \sum_{k \in \mathcal{U}_i^{\text{ut}}} \sqrt{d_i} \mathbf{h}_k \mathbf{s}_i^H + \mathbf{N}. \quad (16)$$

Since \mathbf{s}_i 's are orthogonal to each other, we can extract the signals for group- g as

$$\begin{aligned} \mathbf{y}_g &= \frac{1}{L\sigma_p^2} \mathbf{Y} \mathbf{s}_g = \frac{1}{L\sigma_p^2} \left(\sum_{i=1}^{G^{\text{ut}}} \sum_{k \in \mathcal{U}_i^{\text{ut}}} \sqrt{d_k} \mathbf{h}_k \mathbf{s}_i^H + \mathbf{N} \right) \mathbf{s}_g \\ &= \sum_{l \in \mathcal{U}_g^{\text{ut}}} \sqrt{d_l} \mathbf{h}_l + \frac{1}{L\sigma_p^2} \mathbf{N} \mathbf{s}_g = \sum_{l \in \mathcal{U}_g^{\text{ut}}} \sqrt{d_l} \mathbf{h}_l + \frac{1}{L\sigma_p^2/\sigma_n^2} \bar{\mathbf{n}}_g, \end{aligned} \quad (17)$$

where $\bar{\mathbf{n}}_g \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is the normalized Gaussian white noise vector. Clearly, \mathbf{y}_g only contains the channel vectors of users in group- g while the interference from other groups are completely eliminated by the orthogonal training. The term $\sum_{l \in \mathcal{U}_g^{\text{ut}}} \sqrt{d_l} \mathbf{h}_l$ in (17) is conventionally called as *pilot contamination* [1], which is caused by the same training sequence from different users. Nevertheless, since different channel vectors in group- g have different spatial signatures $\mathcal{B}_k^{\text{ro}}$, we could still extract the individual channel information of each user. For example for user- k in group- g , let us formulate

$$\begin{aligned} \tilde{\mathbf{y}}_{g,k}^{\text{ro}} &= \frac{1}{\sqrt{d_k}} \mathbf{F} \Phi(\phi_k) \mathbf{y}_g = \mathbf{F} \Phi(\phi_k) \mathbf{h}_k + \sum_{l \in \{\mathcal{U}_g^{\text{ut}} \setminus k\}} \sqrt{\frac{d_l}{d_k}} \mathbf{F} \Phi(\phi_k) \mathbf{h}_l + \frac{1}{\sqrt{P_k^{\text{ut}}/\sigma_n^2}} \mathbf{F} \Phi(\phi_k) \bar{\mathbf{n}}_g \\ &= \tilde{\mathbf{h}}_k^{\text{ro}} + \sum_{l \in \{\mathcal{U}_g^{\text{ut}} \setminus k\}} \sqrt{\frac{d_l}{d_k}} \mathbf{F} \Phi(\phi_k) \mathbf{h}_l + \frac{1}{\sqrt{P_k^{\text{ut}}/\sigma_n^2}} \mathbf{F} \Phi(\phi_k) \bar{\mathbf{n}}_g. \end{aligned} \quad (18)$$

From SBEM we know the $\tilde{\mathbf{h}}_k^{\text{ro}}$ can be approximated by $\left[\tilde{\mathbf{h}}_k^{\text{ro}} \right]_{\mathcal{B}_k^{\text{ro}}, \cdot}$, while the estimate of $\left[\tilde{\mathbf{h}}_k^{\text{ro}} \right]_{\mathcal{B}_k^{\text{ro}}, \cdot}$

can be extracted from (18) as

$$\begin{aligned} \widehat{\left[\tilde{\mathbf{h}}_k^{\text{ro}}\right]_{\mathcal{B}_k^{\text{ro}},:}} &= [\tilde{\mathbf{y}}_{g,k}^{\text{ro}}]_{\mathcal{B}_k^{\text{ro}},:} = \frac{1}{\sqrt{d_k}} [\mathbf{I}_M]_{\mathcal{B}_k^{\text{ro}},:} \mathbf{F} \Phi(\phi_k) \mathbf{y}_g \\ &= \left[\tilde{\mathbf{h}}_k^{\text{ro}}\right]_{\mathcal{B}_k^{\text{ro}},:} + \sum_{l \in \{\mathcal{U}_g^{\text{ut}} \setminus k\}} \sqrt{\frac{d_l}{d_k}} [\mathbf{F} \Phi(\phi_k) \mathbf{h}_l]_{\mathcal{B}_k^{\text{ro}},:} + \frac{1}{\sqrt{P_k^{\text{ut}}/\sigma_n^2}} [\mathbf{F} \Phi(\phi_k) \bar{\mathbf{n}}_g]_{\mathcal{B}_k^{\text{ro}},:}. \end{aligned} \quad (19)$$

According to *Property 4* and bearing in mind that $\mathcal{B}_l^{\text{ro}}$ and $\mathcal{B}_k^{\text{ro}}$ are separated at least by one guard interval Ω , we know the entries of $[\mathbf{F} \Phi(\phi_k) \mathbf{h}_l]_{\mathcal{B}_k^{\text{ro}},:}$ in (19) is very small. Then, the estimate of $\tilde{\mathbf{h}}_k^{\text{ro}}$ could be approximated as

$$\hat{\mathbf{h}}_k^{\text{ro}} = \begin{bmatrix} \mathbf{0}^T & \widehat{\left[\tilde{\mathbf{h}}_k^{\text{ro}}\right]_{\mathcal{B}_k^{\text{ro}},:}}^H & \mathbf{0}^T \end{bmatrix}^H, \quad (20)$$

where the two zero vectors $\mathbf{0}$ have appropriate sizes. Hence, the channel estimate of user- k could be computed as

$$\hat{\mathbf{h}}_k = \Phi(\phi_k)^H \mathbf{F}^H \hat{\mathbf{h}}_k^{\text{ro}} = \Phi(\phi_k)^H [\mathbf{F}^H]_{:, \mathcal{B}_k^{\text{ro}}} [\mathbf{I}_M]_{\mathcal{B}_k^{\text{ro}},:} \mathbf{F} \Phi(\phi_k) \mathbf{y}_g. \quad (21)$$

Remark 4: The operation in (19) to get $[\tilde{\mathbf{y}}_g^{\text{ro}}]_{\mathcal{B}_k^{\text{ro}},:}$ from \mathbf{y}_g could be accelerated by taking the partial FFT [23]–[25] rather than a full FFT operation, where the number of complex multiplications to get the τ DFT points is as low as $O(\frac{M}{2} \log_2 \tau)$ when τ is a power of two.

The mean square error (MSE) of $\hat{\mathbf{h}}_k$ in (21) can be expressed as

$$\begin{aligned} \text{MSE}_k^{\text{u}} &= \mathbb{E} \left\{ \left\| \mathbf{h}_k - \hat{\mathbf{h}}_k \right\|^2 \right\} = \mathbb{E} \left\{ \left\| \mathbf{F} \Phi(\phi_k) \mathbf{h}_k - \mathbf{F} \Phi(\phi_k) \hat{\mathbf{h}}_k \right\|^2 \right\} \\ &= \left\| \left[\tilde{\mathbf{h}}_k^{\text{ro}}\right]_{\Xi \setminus \mathcal{B}_k^{\text{ro}},:} \right\|^2 + \left\| \sum_{l \in \mathcal{U}_g^{\text{ut}} \setminus \{k\}} [\mathbf{F} \Phi(\phi_k) \mathbf{h}_l]_{\mathcal{B}_k^{\text{ro}},:} \right\|^2 + \frac{\tau}{P_k^{\text{ut}}/\sigma_n^2}, \end{aligned} \quad (22)$$

where Ξ denotes the complete index set $\{0, \dots, M-1\}$.

It is easy to infer from (22) that the channel estimation error is composed of three parts. The first part is obviously the truncation error from SBEM that only keeps τ DFT points in $\tilde{\mathbf{h}}_k^{\text{ro}}$. The channel leakage from other users in the same group to the current $\mathcal{B}_k^{\text{ro}}$ brings the second error term, which formulates *remaining* pilot contamination. The noise term accounts for the last error part, which is proportional to τ [34]. The truncation error will be reduced if AS of user- k becomes smaller or if the value of τ increases. However, AS is the inherent property of the

environment and τ is generally regulated by the standards, both being most likely uncontrollable. Nevertheless, the inter-user interferences can be reduced when Ω is set to be larger. Hence, if G^{ut} is smaller than τ , then we should reformulate exactly $G^{\text{ut}} = \tau$ groups with evenly distributed users in each group such that the second error term can be reduced.

On the other side, if $G^{\text{ut}} > \tau$, then τ orthogonal training sequences cannot be exclusively assigned to different user groups. In this case, we have to train τ groups after another τ groups until all G^{ut} groups complete the channel estimation period, while the training approach for each τ groups is the same as discussed above.

C. Angle Reciprocity and Downlink Channel Representation

Denote the downlink channel from BS to user- k as $\mathbf{g}_k^H \in \mathbb{C}^{1 \times M}$. Similar to (1), $\mathbf{g}_k \in \mathbb{C}^{M \times 1}$ can be modeled as

$$\mathbf{g}_k = \frac{1}{\sqrt{P}} \sum_{p=1}^P \beta_{kp} \mathbf{a}(\vartheta_{kp}), \quad (23)$$

where $\mathbf{a}(\vartheta_{kp})$ is the steering vector defined in (2) but with different downlink carrier wavelength λ_2 ; ϑ_{kp} is the DOD of the p -th ray that could arrive at user- k , and β_{kp} is the corresponding complex gain. All other parameters have the same definitions as in (1).

Similar to the uplink, we expect that the downlink channel \mathbf{g}_k can also be approximated by SBEM with appropriate spatial signatures $\overline{\mathcal{B}}_k^{\text{ro}}$ and phase shift parameter $\bar{\phi}_k$, i.e.,

$$\mathbf{g}_k = \Phi(\bar{\phi}_k)^H \mathbf{F}^H \tilde{\mathbf{g}}_k^{\text{ro}} \approx \Phi(\bar{\phi}_k)^H [\mathbf{F}^H]_{:, \overline{\mathcal{B}}_k^{\text{ro}}} [\tilde{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}, :} = \sum_{q \in \overline{\mathcal{B}}_k^{\text{ro}}} \tilde{g}_{k,q}^{\text{ro}} \Phi(\bar{\phi}_k)^H \mathbf{f}_q, \quad (24)$$

where $\tilde{g}_{k,q}^{\text{ro}} \triangleq [\tilde{\mathbf{g}}_k^{\text{ro}}]_q$ denotes the q -th element of $\tilde{\mathbf{g}}_k^{\text{ro}} = \mathbf{F} \Phi(\bar{\phi}_k) \mathbf{g}_k$. Following (24), once $\overline{\mathcal{B}}_k^{\text{ro}}$ and $\bar{\phi}_k$ are determined, the estimation of downlink channels \mathbf{g}_k is simplified to estimate the unknown SBEM coefficients $\tilde{g}_{k,q}^{\text{ro}}$.

Since the propagation path of electromagnetic wave is reciprocal, we know that only the signal wave that physically reverses the uplink path can reach the user during the downlink period. Hence, downlink signals that could effectively arrive at the user should have the same DOD spread as the uplink DOA spread, namely, the angle range of ϑ_{kp} 's is the same as the angle

range of θ_{kp} 's. We call this property of the wireless channel as *angle reciprocity*.² For TDD system, the channel gain is also reciprocal such that the overall channel is reciprocal.

Remark 5: Angle reciprocal may not be that useful in conventional MIMO system but is specifically valuable for massive MIMO system where the channel estimation can be decomposed into gain estimation and angle estimation.

Based on the angle reciprocity, $\overline{\mathcal{B}}_k^{\text{ro}}$ can be determined by $\mathcal{B}_k^{\text{ro}}$. Specifically, according to *Property 2*, we will have

$$\sin \theta_{kp} = \frac{q\lambda_1}{Md} = \frac{q'\lambda_2}{Md}, \text{ with } q \in \mathcal{B}_k^{\text{ro}}, q' \in \overline{\mathcal{B}}_k^{\text{ro}}, \quad (25)$$

where λ_1 and λ_2 denote the uplink/downlink carrier wavelengths, respectively. Then the integer set $\overline{\mathcal{B}}_k^{\text{ro}}$ can be expressed as $\overline{\mathcal{B}}_k^{\text{ro}} = \{q'_{\min}, q'_{\min} + 1, \dots, q'_{\max}\}$ with

$$q'_{\min} = \left\lfloor \frac{\lambda_1}{\lambda_2} q_{\min} \right\rfloor, \quad q'_{\max} = \left\lceil \frac{\lambda_1}{\lambda_2} q_{\max} \right\rceil, \quad (26)$$

where $q_{\min} \leq q \leq q_{\max}, \forall q \in \mathcal{B}_k^{\text{ro}}$. Similarly, $\bar{\phi}_k$ can be determined as $\bar{\phi}_k = \frac{\lambda_1}{\lambda_2} \phi_k$.

Remark 6: A key advantage of the proposed antenna array theory based approach over the covariance matrix based method [19], [20] is that the angle of the downlink FDD channel can be predicted from the uplink channel and can be used to simplify the downlink channel estimation.

D. Downlink Training with User Grouping

The key difficulty to apply the conventional downlink channel estimation algorithms for massive MIMO system lies in the requirement that the length of the training has to be no less than the number of antennas as well as the high computational complexity when computing massive parameters. Moreover, the feedback of huge channel state information from user to BS also costs severe overhead.

²If the frequency of the downlink channel is not too far from that of the uplink channel, e.g. less than several GHz, then the reciprocity between DOD and DOA accurately holds. The reason is that for the typical transmission environment, the relative permittivity and the conductivity of the obstacles do not change in the scale of several dozens of GHz [35]. Hence, the reflection and the deflection properties of signals with less than several GHz frequency differences is almost identical in such propagation environment [36], [37].

With SBEM in (24), downlink channel estimation for each user only needs to estimate τ parameters. To reuse the overall τ orthogonal training sequences, let us divide K users into different groups. We first gather users with identical spatial signature $\overline{\mathcal{B}}_k^{\text{ro}}$ into the same cluster. Then, we start to assign clusters into different groups such that the spatial signatures of the clusters in the same group do not overlap and are separated by a certain guard interval, as did in (15). For the ease of exposition, assume that all K users are divided into G^{dt} groups and denote the user index set of the g -th group as $\mathcal{U}_g^{\text{dt}}$.

Let us then take the training of the group- g for example to describe the downlink channel estimation algorithm. Based on SBEM, the effective transmission between τ DFT points in $\overline{\mathcal{B}}_k^{\text{ro}}$ and each user formulates a virtual $\tau \times 1$ MISO downlink system. Hence, to estimate the τ coefficients $[\tilde{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}}$ for each user, we need to transmit τ orthogonal training sequences from the corresponding τ beams \mathbf{f}_q , $q \in \overline{\mathcal{B}}_k^{\text{ro}}$. We then select the orthogonal training matrix for user- k as $\mathbf{S}_k = \varpi_k \mathbf{S}^H \in \mathbb{C}^{\tau \times L}$, where the scalar variable ϖ_k is used to satisfy the transmit power constraint $\text{tr}\{\mathbf{S}_k \mathbf{S}_k^H\} \leq P_k^{\text{dt}}$, and thus $\varpi_k = \sqrt{\frac{P_k^{\text{dt}}}{\tau L \sigma_p^2}}$. Note that, the same orthogonal training matrix $\mathbf{S}^H \in \mathbb{C}^{\tau \times L}$ is reused by different users over their own spatial signatures in the same group.

Then the received signals at user- k of group- g can be expressed as

$$\begin{aligned} \mathbf{y}_k^H &= \mathbf{g}_k^H \left(\sum_{l \in \mathcal{U}_g^{\text{dt}}} \Phi(\bar{\phi}_l)^H [\mathbf{F}^H]_{:, \overline{\mathcal{B}}_l^{\text{ro}}} \mathbf{S}_l \right) + \mathbf{n}_k^H = \sum_{l \in \mathcal{U}_g^{\text{dt}}} \left([\mathbf{F}]_{\overline{\mathcal{B}}_l^{\text{ro}}, :} \Phi(\bar{\phi}_l) \mathbf{g}_k \right)^H \mathbf{S}_l + \mathbf{n}_k^H \\ &= [\tilde{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}, :}^H \mathbf{S}_k + \sum_{l \in \{\mathcal{U}_g^{\text{dt}} \setminus k\}} [\mathbf{F} \Phi(\bar{\phi}_l) \mathbf{g}_k]_{\overline{\mathcal{B}}_l^{\text{ro}}, :}^H \mathbf{S}_l + \mathbf{n}_k^H, \end{aligned} \quad (27)$$

where $\mathbf{n}_k^H \in \mathbb{C}^{1 \times L}$ is the noise vector with its elements distributed as $\mathcal{CN}(0, \sigma_n^2)$. The downlink channel of user- k can be estimated through LS method as

$$[\widehat{\tilde{\mathbf{g}}_k^{\text{ro}}}]_{\overline{\mathcal{B}}_k^{\text{ro}}, :}^H = \mathbf{y}_k^H \mathbf{S}_k^\dagger = [\tilde{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}, :}^H + \sum_{l \in \{\mathcal{U}_g^{\text{dt}} \setminus k\}} \frac{\overline{\varpi}_l}{\varpi_k} [\mathbf{F} \Phi(\bar{\phi}_l) \mathbf{g}_k]_{\overline{\mathcal{B}}_l^{\text{ro}}, :}^H + \mathbf{n}_k^H \mathbf{S}_k^\dagger, \quad (28)$$

where the second term is the interference coming from the DFT points $\mathbf{F} \Phi(\bar{\phi}_l) \mathbf{g}_k$ in set $\overline{\mathcal{B}}_l^{\text{ro}}$, $l \in \mathcal{U}_g^{\text{dt}}, l \neq k$, and is a kind of self-interference.³ Recalling *Property 4*, since $[\mathbf{F} \Phi(\bar{\phi}_l) \mathbf{g}_k]_{\overline{\mathcal{B}}_k^{\text{ro}}, :}$

³There is no pilot contamination in the downlink training, but the reusing of \mathbf{S}^H along different spatial signatures will introduce the self-interference.

contains most channel energy and since $\overline{\mathcal{B}}_l^{\text{ro}}$ is separated at least one guard interval from $\overline{\mathcal{B}}_k^{\text{ro}}$, $[\mathbf{F}\Phi(\bar{\phi}_l)\mathbf{g}_k]_{\overline{\mathcal{B}}_l^{\text{ro}}}$ will be very small compared to $[\tilde{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}}$.

Then the estimate of $\tilde{\mathbf{g}}_k^{\text{ro}}$ can be expressed as

$$\hat{\mathbf{g}}_k^{\text{ro}} = \left[\mathbf{0}^T \widehat{[\tilde{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}}}^H \mathbf{0}^T \right]^H, \quad (29)$$

where the two zero vectors $\mathbf{0}$ have appropriate sizes. Following (24), the channel estimate for user- k in group- g is obtained as

$$\hat{\mathbf{g}}_k = \Phi(\bar{\phi}_k)^H \mathbf{F}^H \hat{\mathbf{g}}_k^{\text{ro}} = \Phi(\bar{\phi}_k)^H [\mathbf{F}^H]_{:, \overline{\mathcal{B}}_k^{\text{ro}}} (\mathbf{S}_k^H)^\dagger \mathbf{y}_k. \quad (30)$$

After channel estimation, each user feeds back τ components $[\hat{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}}$ to BS such that BS can perform the optimal user scheduling and power allocation for the subsequent downlink data transmission. Compared to the feedback of large amount of measurements in [18], the overhead of the newly proposed framework is significantly reduced.

Similar to the analysis of uplink, the downlink MSE of the LS estimator for user- k can be expressed as

$$\begin{aligned} \text{MSE}_k^{\text{d}} &= \mathbb{E} \{ \|\mathbf{g}_k - \hat{\mathbf{g}}_k\|^2 \} = \mathbb{E} \left\{ \left\| \tilde{\mathbf{g}}_k^{\text{ro}} - \hat{\mathbf{g}}_k^{\text{ro}} \right\|^2 \right\} = \left\| \tilde{\mathbf{g}}_k^{\text{ro}} - \left[\mathbf{0}^H \widehat{[\tilde{\mathbf{g}}_k^{\text{ro}}]_{\overline{\mathcal{B}}_k^{\text{ro}}}}^H \mathbf{0}^H \right]^H \right\|^2 \\ &= \left\| [\tilde{\mathbf{g}}_k^{\text{ro}}]_{\Xi \setminus \overline{\mathcal{B}}_k^{\text{ro}}} \right\|^2 + \left\| \sum_{l \in \{\mathcal{U}_g^{\text{dt}} \setminus k\}} \frac{\varpi_l}{\varpi_k} [\mathbf{F}\Phi(\bar{\phi}_l)\mathbf{g}_k]_{\overline{\mathcal{B}}_l^{\text{ro}}}^H \right\|^2 + \frac{\tau^2}{P_k^{\text{dt}}/\sigma_n^2}, \end{aligned} \quad (31)$$

which is also comprised of three parts. The first part is, again, the truncation error from SBEM with only τ spatial signatures. The channel leakage of SBEM from user- k itself results in the second term, which is different from uplink case. The last error term comes from the noise and is proportional to τ^2 , which is consistent to the result in the conventional $\tau \times 1$ MISO downlink system [34]. The self-interference due to pilot reuse will be reduced when the number of users in the same group decreases and the guard interval becomes larger.

Remark 7: From (28), it is seen that user- k does not need the knowledge of spatial signature set $\overline{\mathcal{B}}_k^{\text{ro}}$ and the shift parameter $\bar{\phi}_k$ to perform the estimation of τ channel parameters $\tilde{\mathbf{g}}_k^{\text{ro}}$. This removes the necessity of feedback from BS to the user and is thus a key advantage that make the proposed downlink channel estimation strategy suitable when the mobile users gradually change their positions.

IV. DATA TRANSMISSION WITH USER SCHEDULING

After obtaining the spatial signatures and channel gains of all users, we may schedule users into different groups to enhance the spectral efficiency during the subsequent data transmission period. To achieve this, users in the same group should have non-overlapping spatial signatures such that the inter-user interference could be reduced. Meanwhile, we try to maximize the achievable rate for each group under given power constraint. In the following discussion, we will focus on the downlink case due to page limitation, while the uplink case can be analyzed in a similar manner.

Assume that users are scheduled into G^{dd} groups for data transmission and the user index set of the g -th group is denoted by $\mathcal{U}_g^{\text{dd}}$, $g = 1, 2, \dots, G^{\text{dd}}$. For user- k in the g -th group, its downlink received signal can be expressed as

$$\begin{aligned} y_k^{\text{dl}} &= \sqrt{\rho_k} \mathbf{g}_k^H \mathbf{w}_k x_k + \sum_{l \in \{\mathcal{K}_g \setminus k\}} \sqrt{\rho_l} \mathbf{g}_k^H \mathbf{w}_l x_l + n_k \\ &= \sqrt{\rho_k} (\hat{\mathbf{g}}_k^H + \Delta \mathbf{g}_k^H) \mathbf{w}_k x_k + \sum_{l \in \{\mathcal{K}_g \setminus k\}} \sqrt{\rho_l} (\hat{\mathbf{g}}_k^H + \Delta \mathbf{g}_k^H) \mathbf{w}_l x_l + n_k, \end{aligned} \quad (32)$$

where channel estimation error $\Delta \mathbf{g}_k$ is defined as $\Delta \mathbf{g}_k \triangleq \mathbf{g}_k - \hat{\mathbf{g}}_k$; x_k and \mathbf{w}_k denote the data symbol and the corresponding downlink beamforming vector, respectively; $\rho_k > 0$ denotes the transmit power, and $n_k \sim \mathcal{CN}(0, 1)$ is the received noise.

For downlink massive MIMO systems, the linear matched filter (MF) beamforming was already shown to be a good candidate [38], namely,

$$\mathbf{w}_k = \frac{\hat{\mathbf{g}}_k}{\|\hat{\mathbf{g}}_k\|^2} = \frac{\sum_{q \in \mathcal{B}_k^{\text{ro}}} \hat{g}_{k,q}^{\text{ro}} \Phi(\phi_k)^H \mathbf{f}_q}{\sum_{q \in \mathcal{B}_k^{\text{ro}}} |\hat{g}_{k,q}^{\text{ro}}|^2}. \quad (33)$$

Interestingly, the structure of \mathbf{w}_k indicates the overall beamforming is composed of τ orthogonal sub-beams that correspond to τ spatial signatures of user- k , and the beam gains are exactly selected as their estimated channel gains $\hat{g}_{k,q}^{\text{ro}}$'s. Since users in the same group have non-overlapped spatial signatures, i.e., $\mathcal{B}_k^{\text{ro}} \cap \mathcal{B}_l^{\text{ro}} = \emptyset$, it directly leads to $\mathbf{w}_k^H \mathbf{w}_l = \hat{\mathbf{g}}_k^H \hat{\mathbf{g}}_l = 0$. Note that, the beam orthogonality in our proposed strategy always holds for any finite value of M , as contrast to many other works that require $M \rightarrow \infty$ [19], [39], [40].

Algorithm 1 : User Scheduling Algorithm for Data Transmission

- **Step 1:** Calculate the Euclidean norm of the estimated channel vectors, i.e., $\|\hat{\mathbf{g}}_l\| = \left\| \left[\hat{\mathbf{g}}_l^{\text{ro}} \right]_{\mathcal{B}_l^{\text{ro}},:} \right\|$, for all users.
- **Step 2:** Initialize $g = 1$, $\mathcal{P} = 0$, $\mathcal{U}_g^{\text{dd}} = \emptyset$, $R(\mathcal{U}_g^{\text{dd}}|\mathcal{P}) = 0$, and the remaining user set $\mathcal{U}_r = \{1, \dots, K\}$.
- **Step 3:** For the g -th group, select the user in \mathcal{U}_r with the maximal norm of channel, $l' = \operatorname{argmax}_{l \in \mathcal{U}_r} \|\hat{\mathbf{g}}_l\|$. Set $\mathcal{P} = \rho$, $\mathcal{U}_g^{\text{dd}} = \mathcal{U}_g^{\text{dd}} \cup \{l'\}$ and $\mathcal{U}_r = \mathcal{U}_r \setminus \{l'\}$. Calculate $R(\mathcal{U}_g^{\text{dd}}|\mathcal{P})$ according to (35).
- **Step 4:** Select all those users in \mathcal{U}_r whose spatial signatures are non-overlapping with users in $\mathcal{U}_g^{\text{dd}}$, and denote them by \mathcal{U}'_g , i.e.,

$$\mathcal{U}'_g = \{m \in \mathcal{U}_r \mid \mathcal{B}_m^{\text{ro}} \cap \mathcal{B}_l^{\text{ro}} = \emptyset \text{ and } \operatorname{dist}(\mathcal{B}_m^{\text{ro}}, \mathcal{B}_l^{\text{ro}}) \geq \Omega, \forall l \in \mathcal{U}_g^{\text{dd}}\}.$$

- **Step 5:** If $\mathcal{U}'_g \neq \emptyset$, set $\mathcal{P}' = \mathcal{P} + \rho$, and find a user m' in \mathcal{U}'_g , such that

$$m' = \operatorname{argmax}_{m \in \mathcal{U}'_g} R(\mathcal{U}_g^{\text{dd}} \cup \{m\}|\mathcal{P}').$$

If $R(\mathcal{U}_g^{\text{dd}} \cup \{m\}|\mathcal{P}') \geq R(\mathcal{U}_g^{\text{dd}}|\mathcal{P})$, set $\mathcal{U}_g^{\text{dd}} = \mathcal{U}_g^{\text{dd}} \cup \{m'\}$, $\mathcal{P} = \mathcal{P}'$, $\mathcal{U}_r = \mathcal{U}_r \setminus \{m'\}$ and go to

Step 4; Else, go to **Step 6.**

- **Step 6:** Store $\mathcal{U}_g^{\text{dd}}$ and $R(\mathcal{U}_g^{\text{dd}}|\mathcal{P})$. If $\mathcal{U}_r \neq \emptyset$, let $g = g + 1$, go to **Step 3;** Else, go to **Step 7.**
 - **Step 7:** When the algorithm is stopped, the minimal number of user group G^{dd} is set as the current g , and the optimal user scheduling result is accordingly given by $\mathcal{U}_1, \dots, \mathcal{U}_{G^{\text{dd}}}$.
-

Then (32) can be simplified as

$$y_k^{\text{dl}} = \sqrt{\rho_k} x_k + \Delta \mathbf{g}_k^H \sum_{l \in \mathcal{K}_g} \sqrt{\rho_l} \frac{\hat{\mathbf{g}}_l}{\|\hat{\mathbf{g}}_l\|^2} x_l + n_k. \quad (34)$$

Since we do not assume channel statistics, we can not apply the conventional way [41] to characterize the lower bound of achievable throughput. Nevertheless, in practical case the estimated channel $\hat{\mathbf{g}}_k$ will be used as if it is the true value of \mathbf{g}_k . Hence, we will treat the second term in (34) as gauss noise with the same covariance, which is a worst case and then solve the following

optimization problem to approximately maximize the throughput of each group:

$$\begin{aligned} \max_{\{\rho_k\}} \quad & R(\mathcal{U}_g^{\text{dd}}|\mathcal{P}) \triangleq \sum_{k \in \mathcal{U}_g^{\text{dd}}} \log_2(1 + \rho_k) \\ \text{s.t.} \quad & \sum_{k \in \mathcal{U}_g^{\text{dd}}} \frac{\rho_k}{\left\| \begin{bmatrix} \hat{\mathbf{g}}_k^{\text{ro}} \end{bmatrix}_{\mathcal{B}_k^{\text{ro}},:} \right\|^2} \leq \mathcal{P}, \end{aligned} \quad (35)$$

where \mathcal{P} is the total power constraint for this group. Solutions to (35) can be obtained by the standard water-filling algorithm [42].

We then provide a greedy user scheduling algorithm, where the user with the strongest channel gain will be first scheduled and then the other users with non-overlapping spatial signatures can join the same group only if the achievable sum-rate of the whole group increases afterwards. The detailed steps can be found in Algorithm 1, where the power constraint \mathcal{P} for each group is adjusted dynamically and is proportional to the final number of users in each group.

V. SIMULATIONS

In this section, we demonstrate the effectiveness of the proposed strategy through numerical examples. We select $M = 128$, $d = \lambda/2$, and consider $K = 32$ active users that are gathered into 4 spatially distributed clusters around the mean directions of $\{-48.59^\circ, -14.48^\circ, 14.48^\circ, 48.59^\circ\}$ respectively. The channel vectors of different users are formulated according to (1), where $P = 100$, and α_{kp} is independently taken from $\mathcal{CN}(0, 1)$ for all rays and all users; θ_{kp} is uniformly distributed inside $[\theta_k - \Delta\theta_k, \theta_k + \Delta\theta_k]$, where AS is supposed be $\Delta\theta_k = 2^\circ$ for all users. The system coherence interval is set as $T = 128$ and the default value of τ is assumed to be $\tau = 16$, which is only $1/8$ of the antenna number, and the guard interval for user grouping is set as $\Omega = \tau/4$. The length of pilot L should satisfy $16 \leq L \leq 128$ and is then taken as $L = 16, 32, 64$, respectively. The signal-to-noise ratio (SNR) is defined as $\rho = \sigma_p^2/\sigma_n^2$. The performance metric of the channel estimation is taken as the normalized MSE, i.e.,

$$\text{MSE} \triangleq \frac{\sum_{k=1}^K \left\| \mathbf{h}_k - \hat{\mathbf{h}}_k \right\|^2}{\sum_{k=1}^K \left\| \mathbf{h}_k \right\|^2}.$$

In all examples, the spatial information of users are estimated from the preamble. When $\tau = 16$, the 32 users are scheduled into 2 groups while when $\tau = 8$, the 32 users are scheduled into 4

groups, such that the orthogonal training can be applied to obtain the spatial signatures.

Fig. 6 illustrates the MSE performances of uplink/downlink training in (21) and (30), respectively, as a function of SNR with different training sequence length L . The total power for both uplink and downlink training is constrained to $P_k^{\text{ut}} = P_k^{\text{dt}} = L\rho$ for all users at a given SNR ρ . For the uplink training, $K = 32$ users are divided into $G^{\text{ut}} = 16$ groups. All these 16 groups can be scheduled in the same training length L with $\tau = 16$ available orthogonal training sequences. While for the downlink training, $K = 32$ users are gathered into 4 clusters and are assigned into $G^{\text{dt}} = 1$ group, i.e., they can be scheduled simultaneously in the same training length L too. It is seen from Fig. 6 that when L increases, the MSE performances of uplink/downlink can be improved, since the total training power is proportional to L . It is also seen that as the SNR increases, there is the same error floor for all values of L during both uplink/downlink training. This phenomenon is not unexpected due to the truncation error of SBEM from the real channel and can also be observed in temporal BEM [32]. Meanwhile, since the truncation error is only related with the effective expansion number τ , the error floors will keep the same for different L . Moreover, it can be seen that the uplink MSE performances are generally better than that of downlink for any SNR and L , and each of uplink MSE curves is almost parallel to the corresponding downlink one with a fixed gap between them. This can be inferred by comparing the noise terms of (22) and (31), where the noise power included in the uplink training is only proportional to τ while it is proportional to τ^2 for the downlink training.

Fig. 7 and Fig. 8 compare the proposed channel estimation with the convention LS method for both uplink and downlink cases. To apply the conventional method, $K = 32$ orthogonal training sequences is used for uplink case while 128×128 orthogonal training matrix is used for downlink case. To provide a fair comparison, for any given ρ and L the uplink training power is kept the same $P_k^{\text{ut}} = L\rho$ for both methods, while the total downlink training power for both methods are constrained as $\sum_{k=1}^K P_k^{\text{dt}} = KL\rho$. Note that, to compared with LS, we set $T = 128$, and under such conditions, LS will consume all $T = 128$ symbols for training without any remaining time for data transmission, while the uplink and downlink training overheads of the proposed scheme are only $\lceil \frac{G^{\text{ut}}}{\tau} \rceil L \ll T$ and $G^{\text{dt}} L \ll T$ respectively. It is seen that when

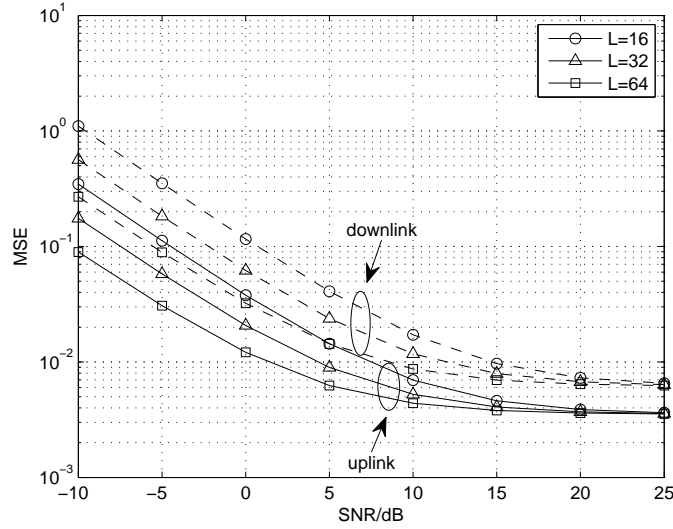


Fig. 6. Comparison of uplink/downlink MSE performances of the proposed SBEM method with $\tau = 16$ and $L = 16, 32, 64$, respectively.

the sufficient number of orthogonal training is available and when the computational complexity is acceptable, then the conventional LS method does not have the error floor for both uplink and downlink cases. Nevertheless, it is interesting to observe that the channel estimation from SBEM outperforms the conventional method when SNR is relatively low. The reasons can be found from (22) and (31) where the proposed method only involves τ components of the noise vector while the conventional LS method includes the whole noise power.

Fig. 9 displays the uplink/downlink MSEs as a function of SNR for $\tau = 8, 16$ respectively with $L = 32$. The total power for both uplink and downlink training is constrained to $L\rho$ for all users at a given SNR ρ . It is seen that as τ increases, the error floors of uplink and downlink MSEs will decline as expected. Interestingly, for downlink training, larger τ will perform better than the smaller τ at higher SNR but will perform worse at lower SNR. The reasons might be inferred from (31) where the three error components are all closely related to the value of $\tau = |\mathcal{B}_k^{\text{ro}}|$. When SNR decreases, the noise component will dominate (31) and thus larger τ will bring more error. While for higher SNR, the truncation error will dominate (31), and thus smaller τ will bring more error. Similar phenomena can also be observed for uplink training,

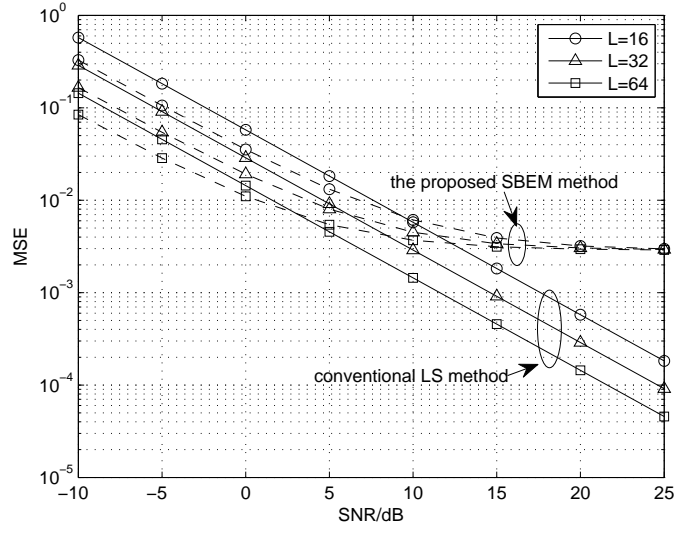


Fig. 7. The uplink MSE performance comparison of the proposed SBEM method and the conventional LS method, with $\tau = 16$ and $L = 16, 32, 64$, respectively.

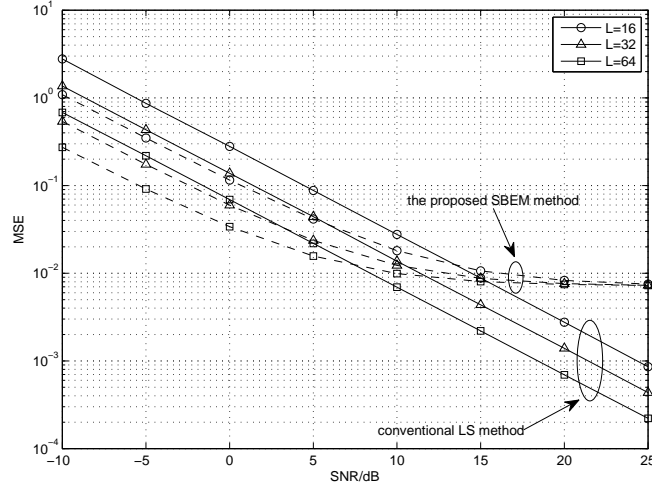


Fig. 8. The downlink MSE performance comparison of the proposed SBEM method and the conventional LS method, with $\tau = 16$ and $L = 16, 32, 64$, respectively.

which can be explained in the same manner over the equation (22).

As spatial rotation (11) is a key technique of our newly proposed strategy, it is then of interest to see how spatial rotation helps improve the performance. We display the channel estimation for

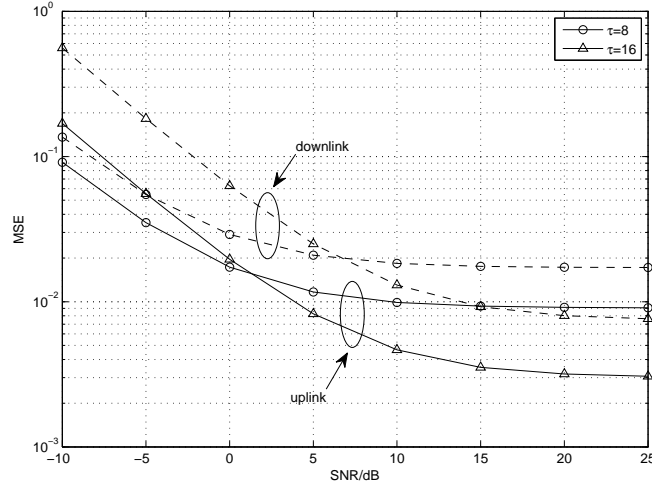


Fig. 9. Comparison of uplink/downlink MSE performances with $L = 32$ and $\tau = 8, 16$, respectively.

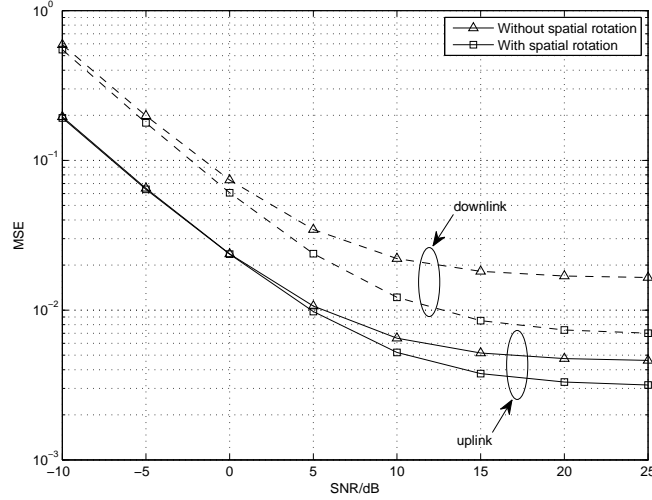


Fig. 10. Comparison of uplink/downlink MSE with and without spatial rotation, where $\tau = 16$ and $L = 32$.

both uplink and downlink with and without spatial rotation in Fig. 10. When the spatial rotation is not adopted, the shift parameter ϕ_k in (11) is set as $\phi_k = 0$ and then the channel estimation procedures remain the same. It is clearly seen from Fig. 10 that the spatial rotation will improve the channel estimation accuracy for both uplink and downlink mainly at high SNRs, i.e., when error floor happens. As explained in *Property 3*, spatial rotation concentrates more channel power

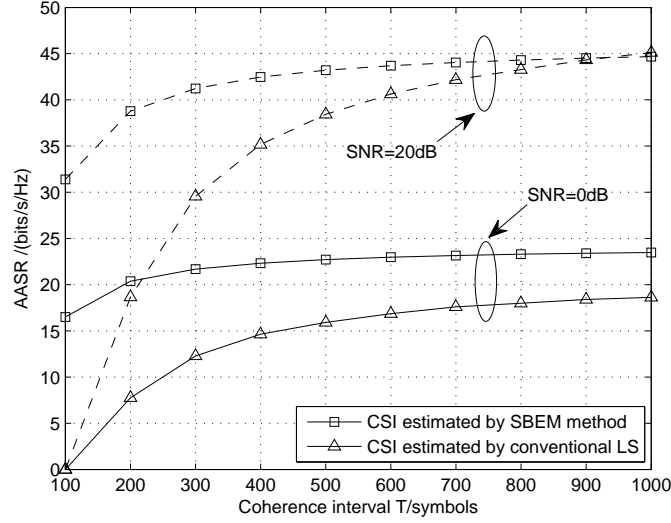


Fig. 11. The average achievable sum rate (AASR) of the proposed SBEM and conventional LS as a function of coherence interval T with $L = 32$.

on fewer DFT points, reduces the power leakage outside $\mathcal{B}_k^{\text{ro}}$, and thus decreases the truncation error of SBEM. Hence, the performance gain by spatial rotation is larger at high SNRs when the truncation error is dominant, while it is not so obvious at the lower SNRs when the noise dominates.

Fig. 11 illustrates the average achievable sum rate (AASR) for the downlink data transmission, defined as

$$\text{AASR} = \left(1 - \frac{T_{\text{pilot}}}{T}\right) \sum_{g=1}^{G^{\text{dd}}} R(\mathcal{U}_g^{\text{dd}} | \mathcal{P}) / G^{\text{dd}}, \quad (36)$$

where T_{pilot} denotes the length of pilot used for training. We take $T_{\text{pilot}} = G^{\text{dt}} L = 32$ for the proposed SBEM and $T_{\text{pilot}} = M = 128$ for the conventional LS. To make the comparison fair, the overall training power and the overall data power within the coherent time T are set as the same for each method. The downlink training procedures for both methods are similar to Fig. 8 with $\tau = 16$ and $L = 32$. Users are scheduled by Algorithm 1. It can be seen from Fig. 11 that the AASR of the proposed SBEM is much higher than that from conventional LS when T is relatively small or when SNR is relatively low. When T becomes large, the training length of LS is small compared to T and then the AASR from conventional LS will approach that from

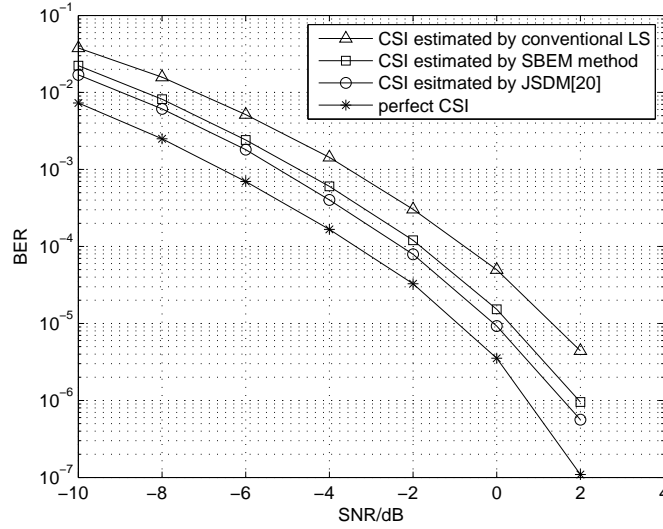


Fig. 12. Comparison of BER performances with perfect CSI, CSI estimated by JSDM [20], CSI estimated by SBEM, and CSI estimated by conventional LS method, respectively, where $\tau = 16$ and $L = 32$.

the the proposed SBEM. Nevertheless, the main disadvantage of the conventional LS method mainly lies in the demand of $L = M$ orthogonal training sequences and the high computational complexity.

Next, we show the bit error rate (BER) performance under QPSK modulation for the downlink data transmission in Fig. 12. Four kinds of CSI are compared, i.e., perfect CSI, CSI from the proposed SBEM method, CSI from the conventional LS method, and CSI from the joint spatial division multiplexing (JSDM) [20]. To keep the comparison fair, the overall training power is set as the same for each method. It is seen that the BER achieved by JSDM is better than the proposed method by about 0.5 dB due to its exploitation of real channel covariances matrices. However, JSDM needs $M \times M$ downlink channel covariance matrices of all users and needs to perform EVD for all these channel covariance matrices to obtain the exact basis vectors of the channels. On the other side, the proposed SBEM utilizes constant Fourier basis vectors for all different users, which eliminates the demand for the downlink channel covariance matrix, and the selection of the spatial signatures can be achieved by efficient FFT. Moreover, both the low-rank methods, SBEM and JSDM, perform better than the LS method at the low SNR region

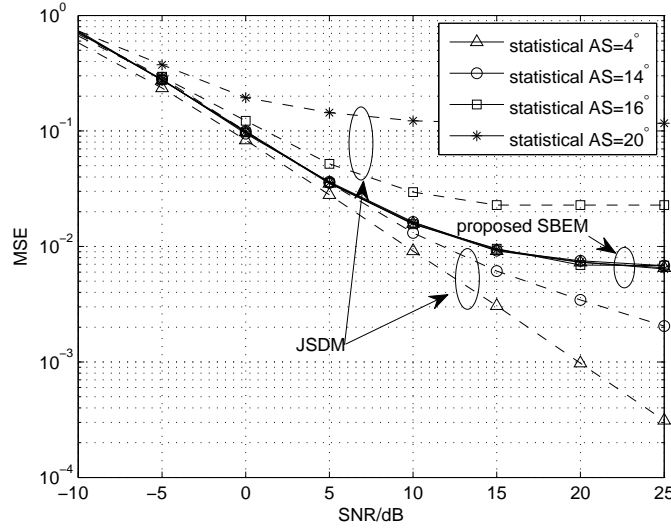


Fig. 13. The DL MSE performance comparison of SBEM and JSDM [20] under the case of user mobility, where $\tau = 16$, $L = 32$.

due to the inclusion of less noise power during channel estimation.

Lastly, we consider the situation with user mobility. Fig. 13 compares the DL MSE performances of the proposed SBEM and the covariance-based JSDM [20], where the instantaneous AS is set as 4° at each time and the statistical AS⁴ is set as $4^\circ, 14^\circ, 16^\circ, 20^\circ$, respectively. It can be seen that as the statistical AS increases, the MSE performances of JSDM deteriorate obviously while the MSE curves from SBEM are not affected so much. The reason lies in that as the statistical AS increases, channel covariances will cover too broad AS and thus are not accurate for the instantaneous channel estimation. Instead, SBEM does not rely on the statistics and thus is more suitable for the mobile user cases.

VI. CONCLUSIONS

In this paper, we investigated the uplink/downlink training and user scheduling for multiuser massive MIMO systems. We exploited the physical characteristics of ULA and proposed a simple DFT-based SBEM to represent the channel vectors with reduced parameter dimensions. It is shown that the basis vectors for different channels formulate the spatial beams towards the

⁴The statistical AS is obtained as user moves around during the measurement period.

users, and thus users could be spatially separated during both the training and data transmission. The conventional headache of pilot contamination is then immediately relieved. Meanwhile, the uplink spatial signatures could also be used for downlink training based on the angle reciprocity of electromagnetic propagation, making the proposed SBEM applicable for both TDD and FDD massive MIMO systems. Channel estimation algorithm for both uplink and downlink were carried out with very few training, and the amount of the feedback could be significantly reduced. To enhance the spectral efficiency, we also proposed a user scheduling algorithm during the data transmission period, where the spatial signatures of users were exploited again. Compared to existing low-rank models, SBEM provides a new simple way to determine the spatial signature information for each user without need of channel statistics and can be efficiently deployed by the FFT operations. Various numerical results were provided to demonstrate the effectiveness and the superiority of the proposed method.

REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [2] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.
- [3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [4] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [5] L. You, X. Gao, X. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, June 2015.
- [6] G. Barriac and U. Madhow, "Space-time communication for OFDM with implicit channel feedback," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3111–3129, Dec. 2004.
- [7] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE J. Sel. Areas Commun.*, no. 31, pp. 192–201, Feb. 2013.
- [8] H. Q. Ngo and E. G. Larsson, "Evd-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. of IEEE ICASSP'12*, Kyoto, Japan, Mar. 2012, pp. 3249–3252.
- [9] R. R. Müller, L. Cottatellucci, M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.
- [10] A. Farhang, A. Aminjavaheri, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Pilot decontamination in CMT-based massive MIMO networks," in *Proc. of IEEE ISWCS'14*, Barcelona, Spain, Aug. 2014, pp. 589–593.
- [11] S. L. H. Nguyen and A. Ghayeb, "Compressive sensing-based channel estimation for massive multiuser MIMO systems," in *Proc. of IEEE WCNC'13*, Shanghai, China, Apr. 2013, pp. 2890–2895.
- [12] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," in *Proc. of IEEE ICC'14*, Sydney, NSW, June 2014, pp. 4626–4631.
- [13] R. C. de Lamare, "Massive MIMO systems: signal processing challenges and future trends," *URSI Radio Science Bulletin*, Oct. 2013.

- [14] S. L. H. Nguyen and A. Ghayeb, "Precoding for multicell massive MIMO systems with compressive rank- q channel approximation," in *Proc. of IEEE PIMRC'13*, London, UK, Sept. 2013, pp. 1227–1232.
- [15] A. Duly, T. Kim, D. Love, and J. Krogmeier, "Closed-loop beam alignment for massive MIMO channel estimation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1439–1442, Aug. 2014.
- [16] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Optimal pilot beam pattern design for massive MIMO systems," in *Proc. of IEEE ACSSC'13*, Pacific Grove, CA, Nov. 2013, pp. 2072–2076.
- [17] P. H. Kuo, H. Kung, and P. A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *Proc. of IEEE WCNC'12*, Shanghai, China, Apr. 2012, pp. 492–497.
- [18] X. Rao and V. K. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, June. 2014.
- [19] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- [20] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.
- [21] C. Sun, X. Gao, S. Jin, M. Matthaiou, Z. Ding, and C. Xiao, "Beam division multiple access transmission for massive MIMO communications," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2170–2184, June 2015.
- [22] W. U. Bajwa, J. Haupt, A. M. Sayeed and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.
- [23] S. He and M. Torkelson, "Computing partial DFT for comb spectrum evaluation," *IEEE Signal Process. Lett.*, vol. 3, no. 6, pp. 173–175, June 1996.
- [24] H. V. Sorensen and C. S. Burrus, "Efficient computation of the DFT with only a subset of input or output points," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1184–1200, Mar. 1993.
- [25] T. Cooklev and A. Nishihara, "Generalized and partial FFT," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 77, no. 9, pp. 1466–1474, Sept. 1994.
- [26] 3GPP, TR 25.996, "Spatial channel model for multiple input multiple output (MIMO) simulations (Rel. 6)," 2003.
- [27] A. Wang, L. Liu, and J. Zhang, "Low complexity direction of arrival (DoA) estimation for 2D massive MIMO systems," in *IEEE GLOBECOM Workshop on Emerging Technologies for LTE-Advanced and Beyond-4G*, Anaheim, CA, Dec. 2012, pp. 703–707.
- [28] Y. Zhu, L. Liu, A. Wang and et al, "DoA estimation and capacity analysis for 2D active massive MIMO systems," in *Proc. of IEEE ICC'13*, Budapest, Hungary, June 2013, pp. 4630–4634.
- [29] L. Liu, Y. Li, and J. Zhang, "DoA estimation and achievable rate analysis for 3D millimeter wave massive MIMO systems," in *Proc. of IEEE SPAWC'14*, Toronto, Canada, June 2014, pp. 6–10.
- [30] A. Hu, T. Lv, H. Gao, et al, "An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 996–1011, Oct. 2014.
- [31] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [32] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct. 1998.
- [33] D. Astely, A. L. Swindlehurst, and B. Ottersten, "Spatial signature estimation for uniform linear arrays with unknown receiver gains and phases," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2128–2138, Aug. 1999.
- [34] M. Biguesh, and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.
- [35] Recommendations, I. T. U. R. "Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900MHz to 100GHz," ITU Recommendations, 2001.
- [36] Y. J. Bultitude, and T. Rautiainen, "IST-4-027756 WINNER II D1. 1.2 V1. 2 WINNER II Channel Models," 2007.
- [37] METIS, Mobile wireless communications Enablers for the Twenty-twenty Information Society, EU 7th Framework Programme project, vol. 6. ICT-317669-METIS.
- [38] G. Dimić and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.
- [39] K. Feng and Y. Liu, "On composite channel estimation in wireless massive MIMO systems", in *Proc. of IEEE GLOCOMW'13*, Atlanta, GA, Dec. 2013, pp. 135–139.

- [40] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1180–1193, July 2014.
- [41] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?", *IEEE J. Sel. Area Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [42] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.